VU VRIJE UNIVERSITEIT AMSTERDAM

VU Research Portal

On the physiology and molecular genetics of salt tolerance in Salicornia

Katschnig, D.

2015

document version
Publisher's PDF, also known as Version of record

Link to publication in VU Research Portal

citation for published version (APA)
Katschnig, D. (2015). *On the physiology and molecular genetics of salt tolerance in Salicornia*. [PhD-Thesis - Research and graduation internal, Vrije Universiteit Amsterdam]. CPI - Koninklijke Wöhrmann.

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:
vuresearchportal.ub@vu.nl

Download date: 05. May. 2023

Chapter 5. Regulation of *SOS1* expression in *Salicornia*: a bioinformatics approach

Diana Katschnig[1], Ken Kraaijeveld[2], Tijs Bliek[3], Jelte Rozema[1], Henk Schat[3]

[1]Systems Ecology, Department of Ecological Science, Faculty of Earth and Life Sciences, VU University Amsterdam, De Boelelaan 1085, 1081 HV Amsterdam, The Netherlands.
[2]Animal Ecology, Department of Ecological Science, Faculty of Earth and Life Sciences, VU University Amsterdam, De Boelelaan 1085, 1081 HV Amsterdam, The Netherlands.
[3]Department of Genetics, Faculty of Earth and Life Sciences, VU University Amsterdam, De Boelelaan 1085, 1081 HV Amsterdam, The Netherlands.

Chapter 5. Regulation of *SOS1* expression in *Salicornia*: a bioinformatics approach

Abstract

The plasma membrane Na^+/H^+-antiporter SOS1 is essential for salt tolerance in plants. In Chapter 4, we observed a ten-fold higher constitutive expression of *SOS1* in the highly salt tolerant salt-accumulating halophyte *Salicornia dolichostachya* in comparison with a lower inducible *SOS1* expression in its glycophytic family member *Spinacia oleracea.* The objective of this study was to investigate the origin of these differences in *SOS1* expression. Copy number variation did not contribute to the differences in expression, since *S. oleracea* and *S. dolichostachya* had both two copies of *SOS1*. Genome walking generated 596bp of the *S. dolichostachya SOS1* promoter, and a *de novo* draft genome assembly of the diploid *Salicornia brachystachya* generated 2372bp of the *SbSOS1* promoter. Analysis of the *cis*-regulatory DNA elements in these promoters in comparison with the *Arabidopsis thaliana SOS1* promoter showed that the *Salicornia SOS1* promoters contained fewer *cis*-regulatory elements with an ABRE-variant than the *AtSOS1* promoter. Moreover, the DRE/CRT element was uniquely present in the *AtSOS1* promoter. These elements are the binding sites for stress-responsive transcription factors, and might be responsible for the induction of *AtSOS1* in response to salt stress. Absence of these elements in the *SbSOS1* promoter might explain the lack of *SOS1* inducibility in *Salicornia* in response to salinity, but not the high-level expression of *SdSOS1*.

5.1. Introduction

Salinity causes considerable growth-reductions in the majority of plant species. Over the last decades substantial effort has been dedicated to improve salt tolerance in crops (Flowers and Flowers 2005, Munns and Tester 2008). Some plant species are naturally tolerant to salinity. These halophytes occur widely spread over different plant families and orders (Flowers *et al.* 2010), which makes it likely that differences in salinity tolerance between and among halophytes and glycophytes result from differences in level and location of gene expression, rather than from genes that are unique to halophytes (Kant *et al.* 2006; Rozema and Schat 2013). To improve salt tolerance in plants, identification of genes that contribute to the salt tolerance in halophytes is necessary, but still complicated due to an overall lack of genetically accessible halophytes thus far.

To date, *Eutrema halophilum* (= *Thellungiella halophila*) is the only halophyte that is genetically accessible. *E. halophilum* has been chosen as a halophyte genetic model because of its high genetic similarity with *Arabidopsis thaliana* (92% gDNA identity), and the possibility to genetically transform *E. halophilum* by the flower dip protocol (Inan *et al.* 2004). However, *E. halophilum*, like *A. thaliana*, has the drawback that it is a multiple stress tolerator. Moreover, its salt tolerance is based on 'endurance' of periods of high salinity and avoidance of salt accumulation, rather than on its ability to maintain growth under high levels of salt exposure. On the contrary, its growth capacity is strongly reduced already at relatively low salinity levels, in comparison with many coastal halophytes (Flowers and Colmer 2008). Therefore, in addition to *E. halophilum*, it would also be interesting to study more extreme halophytes, such as those belonging to the family of Amaranthaceae, who combine high levels of salt tolerance with high salt accumulation in the shoot (Flowers *et al.* 2010). Unfortunately, the halophytes of the Amaranthacea family are, as yet, hardly genetically accessible, which hampers the research into the salt tolerance mechanisms of these species.

The species of the genus *Salicornia* are among the most salt-tolerant within the Amaranthacea. They occur globally (except Australia and South America) in salt marshes (Kadereit *et al.* 2007) and do not have salt glands or salt bladders. *Salicornia* species use Na^+ as a energetically cheap osmolyte to adjust to the low external water potential created by a highly saline environment. They can accumulate NaCl up to 50% of their dry weight without growth reduction, and their growth optimum lies around 200-300 mM NaCl (Katschnig *et al.* 2013). *Salicornia* species are being studied because of their exceptionally high salt tolerance

and their potential as vegetable and oil-seed crops, however, to date they lack 'genetics' (Flowers and Colmer 2008). Genetic accessibility of *Salicornia* would disclose valuable information to identify traits that underlie their exceptionally high tolerance to salinity.

In an attempt to unravel the salt-tolerance mechanisms of *Salicornia*, we looked previously at the differences in gene expression between *Salicornia dolichostachya* and its glycophytic family member *Spinacia oleracea* (Chapter 4). We found that the *SOS1* gene was ten-fold higher expressed in *S. dolichostachya* compared with *S. oleracea*, and the expression in *S. dolichostachya* was constitutive, whereas in *S. oleracea* leaves *SOS1* expression was upregulated in response to the salt treatment. *SOS1* encodes a Na^+/H^+-antiporter that is located at the plasma membrane, which transports Na^+ against the electrochemical gradient from the cytoplasm into the apoplast (Oh *et al.* 2010). In *Arabidopsis thaliana*, *SOS1* is expressed in the Xylem Parenchyma Cells (XPCs) and in the epidermal cells of the root tips (Oh *et al.* 2009). In the XPCs, it functions in Na^+ transport into the xylem, which increases the Na^+ concentrations in the shoot. In the root epidermal cells, it functions in excretion of Na^+ from the roots into the external environment, thereby lowering the Na^+ concentrations in the roots. So, the effect of SOS1 on Na^+ concentrations inside the plant is expected to depend on both the level and the location of expression. Differences in the expression patterns of genes can be caused by both copy-number expansion or by differences in regulatory elements acting on the gene (Hanikenne *et al.* 2008).

SOS2 and SOS3 are involved in the signaling pathway leading to SOS1 activation. SOS3 is a mirystoylated calcineurin-like calcium-binding protein (Ishitani *et al.* 2000). Increased cytoplasmic Ca^{2+} activates SOS3, which binds the serine/threonine protein kinase SOS2 (Halfter *et al.* 2000). The activated SOS2-SOS3 complex activates SOS1 by phosphorylation (Qui *et al.* 2002). Other Na^+ transporters, such as NHX1 and Ca^{2+}/H^+ exchanger (CAX), might also be regulated by this complex (Cheng *et al.* 2004; Qui *et al.* 2004). Furthermore, it has been shown that the salt stress induced upregulation of *SOS1* observed in *A. thaliana* is partly controlled by the SOS2-SOS3 complex (Shi *et al.* 2000).

To unravel the causes of the differences in expression of *SOS1* between *S. oleracea* and *S. dolichostachya*, we determined the copy number of the *SdSOS1* gene and screened for putative *cis*-regulatory elements in the *SdSOS1* promoter region, in comparison with the *AtSOS1* promoter. Since we could only obtain a small fraction of the promoter sequence of *S. dolichostachya* with classical genome walking methods, we decided to use high throughput parallel sequencing technologies to obtain a longer fragment of the *SOS1* promoter sequence. Because *de novo* genome assembly would become notoriously complicated with the

tetraploid *S. dolichostachya*, we made a *de novo* draft genome-assembly of another equally salt-tolerant, but diploid, *Salicornia* species: *Salicornia brachystachya* (Huiskes *et al.* 1985). For *S. brachystachya,* we also determined *SOS1* copy number and screened for putative *cis*-regulatory elements in the *SOS1* promoter region.

5.2. Materials and Methods

5.2.1. Plant material and gDNA extraction

In October 2012, seeds of the tetraploid *Salicornia dolichostachya* Moss were collected below the mean high water line (MHWL) in a coastal salt marsh along the Wadden Sea at Lutjestrand in Wieringen, The Netherlands. In October 2013, seeds of the diploid *Salicornia brachystachya* were collected above the MHWL in a coastal salt marsh (De Slufter, Texel), The Netherlands. *Spinacia oleracea* L. seeds were obtained from a commercial supplier (Tuin plus Service, Holland). Seeds were sown on peat soil (Seed pot soil; Jongkind, Aalsmeer) and grown for 60 days (*S. dolichostachya* and *S. brachystachya*) or 12 days (*S. oleracea*). Genomic DNA (gDNA) was extracted from the shoot of one single plant following a cetyltrimethylammonium bromide (CTAB)-based protocol (Rogers and Bendich 1985). The integrity of the gDNA was checked on 1% agarose gel and gDNA quantity was determined using NanoDrop spectrophotometry (NanoDrop^{TM} 1000, Thermo Scientific).

5.2.2. Determination of *SOS1* gene copy number in *S. dolichostachya* and *S. oleracea* using Southern Blot

5.2.2.1. Preparation of the blot

Twenty μg of gDNA from *S. dolichostachya* and *S. oleracea* was cut with six different restriction enzymes with a six-base restriction site: BAM II, Eco RI, Eco RV, Hind III, Pst I and Xba I. The restriction reactions were carried out o/n at 37 °C in separate vials in a total volume of 100 μl containing; 20 μg gDNA, 10 μl of the appropriate buffer, 20 unit of enzyme and, if required, BSA. Thereafter, products were cleaned and concentrated and run on a gel overnight. The gel was checked for the expected smear of bands, and subsequently gently shaked with 0.25 M HCl, denaturating solution (1.5M NaCl/0.5 M NaOH), renaturating solution (1.5M NaCl/0.5M Tris-HCL pH7.0) and 10x SSC. The gel was blotted on a nitrocellulose filter by capillary rise.

5.2.2.2. Probe design and hybridization for Southern Blotting

Gene specific primers for *SdSOS1* and *SoSOS1* on gDNA were designed for the construction of the labelled probes, *SdSOS1* forward 5'-GTTTTGCTTGCTGGACCAGGTG and reverse 5'-GAGCTACTACAGCTACAGGATC, *SoSOS1* forward 5'-GTTATACTTGCTGGACCAGGAG and reverse 5'-CTTAGAAGTCCACCGAGCAAC. DNA probes were labelled with ^{32}P-dUTP which was incorporated by PCR. 1 ng DNA template was used for each probe. The probes were hybridized to the membranes overnight by rotating at 65 °C in hybridization buffer. After hybridization, the membranes were washed at 65°C in 2 x SSC and 0.1% (w/v) SDS, followed by a more stringent washing step. Thereafter, membranes were sealed in plastic and stored up to 72 hours. The membrane was exposed to a phosphorus screen, and this screen was subsequently scanned with a Storm 820 Phospho-imager.

5.2.3. Obtaining the *S. dolichostachya* promoter region by genome walking

Described in Chapter 4.

5.2.4. Illumina HiSeq 2000 sequencing and *de novo* genome assembly *Salicornia brachystachya*

5.2.4.1. DNA library preparation and Illumina sequencing

Because we could not obtain a long promoter fragment of the *SOS1* gene from the tetraploid *S. dolichostachya*, we decided to sequence the genome of diploid *Salicornia brachystachya.* Two µg of gDNA of *S. brachystachya* was sent for library construction and paired-end sequencing to BGI Tech Solutions, Hong Kong. The gDNA was sheared to a fragment size of approximately 500 bp. The DNA library was generated by BGI Tech Solutions using the library construction protocol for 500 bp short-insert libraries (Illumina Inc.). The library was sequenced with a 101bp paired-end sequencing strategy on a single lane of the Illumina HiSeq 2000 (Illumina Inc). The raw sequence reads were quality filtered by removing the adapter sequences and discarding low-quality reads. After quality filtering, we obtained approximately 36.6 Gb of clean 100 bp paired-end reads. We used FastQC to perform quality-control checks on the sequence data.

5.2.4.2. De novo genome-assembly of S. brachystachya

De novo draft genome assembly of the quality-filtered reads was performed with Platanus v1.2.1 (Kajitani *et al.* 2014). Platanus uses a de Bruijn graph and Eulerian path-based assembly approach, and paired-end reads for scaffolding of contigs. Platanus automatically optimizes k-mer sizes in constructing the de Bruijn graphs. Assemblies were performed on the Dutch national e-infrastructure with the support of SURF Foundation. We estimated the output quality of the draft assembly by the amount of contigs larger than 200 bp, the length of the longest contig, the total length of the assembled contigs and the N50 scaffold length, which refers to weighted median of the total length of assembled scaffolds.

5.2.5. Determination of *SOS1* gene copy-number in *S. brachystachya* using read depth

To determine copy number of the *SOS1* gene in *S. brachystachya,* the *de novo* draft genome assembly was searched (tblastx) with the *SdSOS1* gene sequence. To evaluate if the assembler incorrectly merged different copies of *SbSOS1*, paired-end quality-filtered reads were aligned with Bowtie2 v2.2.3 against the *de novo* assembled draft genome-sequence that we generated with Platanus. Only the best alignment was recorded. The output SAM-alignment file was converted into a BAM file and sorted using SAMtools v1.0. The protein

sequences of *SOS1* and 15 randomly chosen genes from the Core Eukaryotic Gene mapping Approach (CEGMA) database (At1g08780, At1g08830, At1g10490, At1g11870, At1g11890, At1g 12230, At1g12840, At1g22780, At1g25350, At1g34130, At1g36240, At1g44900, At1g48830, At1g49910) were mapped to the *S. brachystachya* draft genome using tblastn with an threshold e-value <1e-5. The CEGMA genes are highly conserved and of low copy-number, and they are expected to be present in all eukaryotes (Parra *et al.* 2007). BEDtools v2.20.1 was used on the BAM file containing the read alignment to calculate average read-depths of the regions where *SOS1* and the CEGMA genes were mapped.

5.2.6. Obtaining the *S. brachystachya* promoter region using our *de novo* draft genome assembly

To obtain the *S. brachystachya SOS1* promoter sequence, the *S. dolichostachya SOS1* gene sequence (EMBL accession number HG799054) was mapped to the *S. brachystachya de novo* draft genome using tblastx (e-value <1e-5). We recorded the sequence located immediately upstream of the start of the *SOS1*-gene sequence and refer to this region as the *SbSOS1* promoter sequence. We designed several primers on this promoter region (Table S2) and a reverse primer on the 5'region of the *SOS1* gene to check for proper genome assembly of this fragment. The obtained fragments of the *SbSOS1* promoter were cloned in the pGEM-T Easy vector (pGEM®-T Easy Vectors Systems, Promega Madison, WI, USA) using the manufacturers protocol, and after isolation prepared for Sanger sequencing using the Big Dye terminator kit (Applied Biosystems, Life Technologies Corporation, Carlsbad, CA, USA). The *SbSOS1* promoter sequence has been deposited in the European Nucleotide Archive under the accession number LN680868.

5.2.7. Putative *cis*-acting elements in the *SdSOS1* and *SbSOS1* promoter sequences

To find potential indications for differential activity of the *SOS1* promoters in *S. dolichostachya* and *S. brachystachya*, in comparison with the *SOS1* promoter of the glycophyte *Arabidopsis thaliana*, we identified putative *cis*-acting regulatory DNA elements in the promoter sequence of the *SOS1* gene. To identify these elements, we used the PLACE program (Higo *et al.* 1999). Promoter sequences were aligned with ClustalW2 (Fig. S2) and putative *cis*-acting regulatory DNA elements were identified (Table S2).

5.3. Results

5.3.1. Both *Salicornia dolichostachya* and *Spinacia oleracea* have two copies of the *SOS1* gene

Because we observed considerable differences in expression of the *SOS1* orthologous gene between the *Salicornia dolichostachya* and *Spinacia oleracea* (Chapter 4) we questioned whether this was due to a higher copy number of *SOS1* in *S. dolichostachya* or to regulatory differences in gene expression. Southern blot analysis showed that both species have two copies of *SOS1* (Fig. 1).

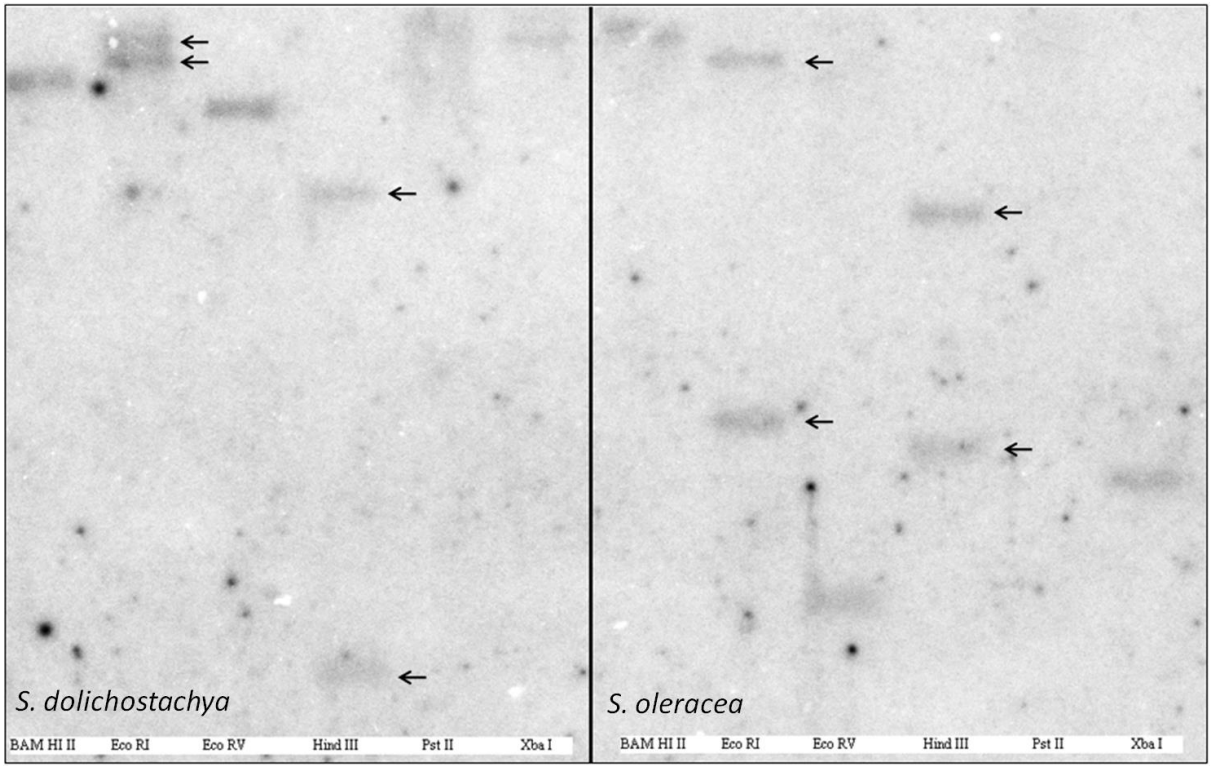

Figure 1. Southern blot of *Salicornia dolichostachya* (left) and *Spinacia oleracea* (right). DNA of both species was cut with six restriction enzymes (in lanes from left to right: BAM HI II, Eco RI, Eco RV, Hind III, Pst II, Xba I). Arrows indicate the two copies of the *SOS1* gene.

5.3.2. Genome walking resulted in a 596bp fragment of the *SdSOS1* promoter

Genome walking generated 596bp immediately upstream of the 5'-end of *SdSOS1* (Chapter 4). This fragment is referred to as the partial *SdSOS1* promoter.

5.3.3. *De novo* draft genome assembly of *Salicornia brachystachya*

5.3.3.1. Illumina sequencing

After sequencing of 500bp insert fragments of *S. brachystachya* gDNA on a single lane of a flow cell of the Illumina HiSeq 2000, we obtained approximately 36.6Gb of clean sequence reads (Table 1). The genome size (1C) of *Salicornia europaea* (=*Salicornia brachystachya*) is expected to be 1.35kMb (Plant DNA c-value database release 6.0), which would imply that we obtained about 27.4x coverage of the genome. However, our analysis of read mapping of the CEGMA genes indicated a coverage of approximately 65x times, which would imply that the genome size of *S. brachystachya* is around 0.56 kMb, and less than half the size as registered in the Plant DNA c-value database. We performed quality control checks on the quality-filtered reads with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The parameters of the quality control check (Table 1) indicated that the sequenced data was of sufficient quality to perform genome assembly.

Table 1. Summary of the production of genomic DNA sequences after quality filtering. DNA was extracted from the shoot of *Salicornia brachystachya* and sequences were produced with the Illumina HiSeq 2000, based on a 101bp paired-end sequence strategy with a 500bp short insert library.

Insert size (bp)	Read length (bp)	# Clean reads	# Clean bases	Phred quality score	GC (%)
500	100	366246776	36624677600	38	35.2

5.3.3.2. De novo genome-assembly with Platanus

De novo genome assembly of *S. brachystachya* with the short-read assembly program Platanus generated in total 311751 contigs with a length of 200bp or longer (Table 2). 50% Percent of the assembled draft genome was situated in blocks of 2212bp (the Scaffold N50 size) or longer, and the longest contig had a length of 29951bp. The total length of assembled contigs was close to $20*10^6$, which would imply—based on a genome size of 0.56 kMb—that we assembled approximately 36% of the total genome.

Table 2. Output quality of the *de novo* draft genome assembly of *Salicornia brachystachya* generated with Platanus v1.2.1. The following criteria were used to evaluate the quality of the genome assembly: contigs >200bp = the amount of contigs larger than 200bp, the length of the longest contig, the total length of the assembled contigs and the scaffold N50 score, which refers to the weighted median of the total length of assembled scaffolds.

de novo assembler	Contigs >200bp	Longest contig (bp)	Total length assembled contigs	Scaffold N50
Platanus	311751	29951	195749480	2212

5.3.4. Determination of *SOS1* gene copy-number in *S. brachystachya* using read depth

We searched the *de novo* draft genome assembly of *S. brachystachya* for copies of the *SbSOS1* gene. We could identify the complete *SbSOS1* gene. It was fragmented and located on seven different scaffolds. To check for incorrectly merged copies of *SbSOS1*, we used read depth. It can be assumed that the probability of being sequenced is equal for every base in the genome, therefore, the read depth, which is the number of reads mapped to a particular sequence in the genome, is proportional to the number of times this sequence appears in the genome. The *SbSOS1* gene had an average read depth of approximately 65x. The read depth of 15 randomly chosen CEGMA genes, which are expected to be highly conserved and present in low copy-number, was also approximately 65x. This implied that the *SbSOS1* gene did not get incorrectly merged by the assembler and that *S. brachystachya* had one copy of *SOS1*.

5.3.5. Obtaining the *S. brachystachya* promoter region using the *de novo* draft genome assembly

We searched the *de novo* draft genome assembly of *S. brachystachya* with the *SdSOS1* gene sequence, and recorded the 2372bp sequence located immediately upstream of the start of the 5'-end of the *SbSOS1* gene. We refer to this region as the *SbSOS1* promoter (Fig. S1). We evaluated the correctness of the assembly by designing primers on this *SbSOS1* promoter as well as downstream of the ATG translation initiation codon of the *SbSOS1* gene and subsequently sequencing of the fragments with a Sanger-sequencing technology. The sequence that we obtained with Sanger sequencing was similar to our *de novo* generated *SbSOS1* promoter, which confirmed a correct assembly of the *SbSOS1* promoter by Platanus.

5.3.6. Putative *cis*-acting elements in the *AtSOS1* and *Salicornia SOS1* promoter

We aligned the *SbSOS1* promoter, the partial *SdSOS1* promoter and the *AtSOS1* promoter (Fig. S2). The *SbSOS1* promoter and the partial *SdSOS1* promoter shared 98.6% similarity. The *AtSOS1* promoter and the *SbSOS1* and partial *SdSOS1* promoter shared, respectively, 50.4% and 45.6% similarity. Because the three promoters differed in length, we split the analysis of the promoter sequences for putative *cis*-regulatory elements in three parts corresponding to the length of the consecutive shortest sequence based on the alignment. In the partial *SdSOS1* promoter roughly the same elements were present as in the corresponding part of the *SbSOS1* promoter (Table S2). The putative *cis*-regulatory elements in *Sd* and *Sb* were located at the same positions. The *AtSOS1* promoter had 13 unique *cis*-regulatory elements in comparison with the *Salicornia SOS1* promoter (Table S2). Among these unique elements in the *AtSOS1* promoter were the ABRE-related sequence, ABRE G-BOX, CBF, DRE/CRT and several light responsive *cis*-regulatory elements (Table 3). The *SbSOS1* promoter possessed 44 different putative *cis*-regulatory DNA elements that were unique for *SbSOS1* promoter and did not occur in the *AtSOS1* promoter (Table S2). Among the unique *cis*-regulatory elements in the *SbSOS1* promoter were a MYC-like sequence and the palindromic MYC-binding site, several auxin responsive *cis*-regulatory elements and a TAAAG-motif (Table 3). The ARR1-binding element was present in considerably higher numbers in the *SbSOS1* promoter compared with the *AtSOS1* promoter (Table 3 and Table S2).

Table 3. Potential *cis*-acting regulatory DNA elements present in the *Arabidopsis thaliana SOS1* promoter and the *Salicornia brachystachya SOS1* promoter. A plus or minus-sign behind the location indicates the plus-strand and minus-strand, respectively. Identification of potential *cis*-acting regulatory elements was conducted with PLACE (Higo *et al*. 1999).

Species	*Cis*-regulatory element	Sequence	Location (nucleotides upstream of ATG translation initiation codon)	Function	Reference
Uniquely present in *A. thaliana*	ABRE-related sequence	(C/A)ACG(C/T)G(T/C/G)	116+, 116-	Abscisic acid (ABA) responsive element. $Ca2^{+}$ responsive *cis*-regulatory element.	Yamaguchi-Shinozaki *et al.* 2005 Kaplan *et al.* 2006
Uniquely present in *A. thaliana*	G-Box ABRE	CACGTG	115+, 115-	G-box, involved in light-responsiveness. Binding site of G-Box binding-factor (GBF).	Menkens *et al.* 1995 Lopez-Ochoa *et al*. 2007 Priest *et al.* 2009
Uniquely present in *A. thaliana*	CBF	(G/A)(C/T)CGAC	220-	Binding site for C-repeat binding factors (CBFs), also known as dehydration responsive element (DRE) binding proteins (DREBs).	Yamaguchi-Shinozaki *et al.* 2005
Uniquely present in *A. thaliana*	DRE/CRT	(G/A)CCGAC	220-	Dehydration-Responsive Element/C-repeat. Involved in drought, dehydration and cold stress.	Yamaguchi-Shinozaki *et al.* 2005
Uniquely present in *A. thaliana*	LTRE	CCGAC	220-	Low-temperature responsive-element.	Yamaguchi-Shinozaki *et al.* 2005
Uniquely present in *A. thaliana*	PRE	(C/G)CGA(C/T)N(A/G)N NNNNNNNNNNNNN (A/T/C)(A/T/G)	239-	Consensus sequence of plastid response element (PRE), a *cis*-regulatory element involved in light responsiveness.	Von Gromoff *et al.* 2006
Uniquely present in *A. thaliana*	rbcS consensus	AATCCAA	233+, 732+	Consensus sequence of ribulose-1,5-biphophate carboxylase. *Cis*-regulatory element involved in light responsiveness.	Menkens *et al.* 1995 Lopez-Ochoa *et al.* 2007
More in *S. brachystachya*	ARR1-binding element	NGATT	42-, 109-, 144+, 285-, 584-, 787-, 862+, 876+, 905+, 966+, 990+, 1125+, 1142+, 1354+, 1511-, 1601+, 1624+, 1729+, 1779+, 1848+, 1897-, 2152-, 2318+, 2371+	ARR1-binding element, a cytokinin-responsive binding-element.	Sakai *et al.* 2001
Uniquely present in *S. brachystachya*	Activation sequence factor-1	TGACG	733+, 250+	Activation sequence factor-1 (Asf-1), a *cis*-regulatory element involved in auxin and/or salicylic acid responsiveness.	Garretón *et al.* 2002
Uniquely present in *S. brachystachya*	SAUR	CATATG	363-, 363+, 573-, 573+, 1733+, 1733-	Sequence in the NdeI restriction endonuclease site element (NDE), which is part of the small auxin up RNA (SAUR), a *cis*-regulatory element involved in auxin responsiveness.	Li *et al.* 1994
Uniquely present in *S. brachystachya*	MYC-like sequence	CATGTG	2202+	Binding site for MYC, a *cis*-regulatory element involved in responsiveness to drought/dehydration and abscisic acid induction.	Abe *et al.* 2003 Tran *et al.* 2004
	MYC-binding site	CACATG	2202-		
Uniquely present in *S. brachystachya*	SORLIP	GCCAC	10+, 244-, 1959+, 1959-, 2206-	A *cis*-regulatory element over-represented in light-induced promoters.	Hudson *et al.* 2003
Uniquely present in *S. brachystachya*	SuRE	GAGAC	385-, 1801+	Sulphur-responsive element, a *cis*-regulatory element involved in auxin responsiveness.	Maruyama-Nakashita *et al.* 2005
Uniquely present in *S. brachystachya*	TAAAG-motif	TAAAG	93-, 935-, 1106-, 1471-, 1877-, 2054-, 2254+	Target site for trans-acting Dof1 protein controlling guard cell-specific gene expression.	Plesch *et al.* 2002

5.4. Discussion

Because unique salt-tolerance genes do not appear to exist, salt tolerance in halophytes is presumed to arise from variations in gene expression (Rozema and Schat 2013). Recent studies in metallophytes have demonstrated that plant adaptation to metal-toxic soils, particularly in species adopting the "hyper-accumulator strategy" (Baker 1981), is associated with constitutive expression and genomic copy number expansion of the, usually metal transporter-encoding, tolerance genes (Hanikenne *et al.* 2008; Shahzad *et al.* 2010; O Lochlainn *et al.* 2011). To check whether copy number expansion also plays a role in the previously observed large difference in *SOS1* expression between *Spinacia oleracea* and *Salicornia dolichostachya* (Chapter 4), we compared the genomic copy numbers with Southern blotting. Both species appeared to have two copies of the *SOS1* gene (Fig. 1). Moreover, based on the read-depth analysis, *Salicornia brachystachya* seems to have only one copy of the *SOS1* gene. That both *Salicornia* species in comparison with *S. oleracea* had the same number or fewer copies of *SOS1* suggests that the previously observed difference in *SOS1* expression between *S. oleracea* and *S. dolichostachya* were not due to copy number variation and therefore must have a different origin.

Cis-regulatory elements in the promoter sequence play a key role in the regulation of gene expression. In general, most *cis*-regulatory elements are located within 2000bp upstream of the ATG translation initiation codon, although occasionally the regulatory elements are located even further upstream or might be present in introns. The *A. thaliana SOS1* promoter sequence is fairly short: 981bp (The Arabidopsis Information Resource). The fragment of the *S. dolichostachya SOS1* promoter that we obtained with genome walking was only 596bp (Chapter 4). Therefore, we tried to obtain a longer fragment of the *S. brachystachya SOS1* promoter via Illumina short-read sequencing and subsequent genome-assembly, which generated 2372bp of the *SbSOS1* promoter (Fig. S1). This result showed that applicability of short-read sequencing is a fruitful strategy to obtain previously unknown sequences. Among the differences in *cis*-regulatory elements in the *SOS1* promoters of the species (Table S2), we selected a few elements that might play a role in the observed differences in *SOS1* expression between *S. dolichostachya* and *S. oleracea* (Table 3). However, it is context-dependent if these *cis*-regulatory elements are functional (Priest *et al.* 2009), and the role of these potential *cis*-regulatory elements in the promoters of *AtSOS1* and *Salicornia SOS1* remains to be experimentally demonstrated.

Among the unique *cis*-regulatory elements in the *AtSOS1* promoter sequence were: the Abscisic Acid Responsive Element related sequence (ABRE-related element), the ABRE G-box, the Binding site for *cis*-repeat Binding Factors (CBF), the Dehydration Responsive Element/C-Repeat (DRE/CRT) and Low Temperature Responsive Element (LTRE) (Table 3 and Table S2). Abscisic acid (ABA) is a major plant hormone. Its synthesis is induced by a variety of abiotic stresses, and it accumulates during dehydration (Davies 2010). Many osmotically responsive genes are induced by ABA (Mizoi *et al.* 2012). ABA binds to *cis*-regulatory DNA elements containing a variant of the ABRE-sequence (Yamaguchi-Shinozaki *et al.* 2005). The ABRE-related sequence ((C/A)ACG(C/T)G(T/C/G)) is such an ABRE-sequence. It is found in putative promoter-sequences of Ca^{2+}-responsive upregulated genes, and localized between 100 and 500bp upstream of the ATG translation initiation codon (Kaplan *et al.* 2006). Intracellular Ca^{2+} is an important signal in many stress-activated pathways, and thought to act as a fast second-messenger in response to ABA (Hirayama and Shinozaki 2007). In the *AtSOS1* promoter the ABRE-related sequence is indeed located in this area, 461bp upstream of the ATG translation initiation codon. Both *Salicornia SOS1* promoters and the *AtSOS1* promoter contained two or more classical ABRE-motifs (ACGTG), and the *AtSOS1* promoter additionally contained the related ABRE G-Box (CACGTG). Two classical ABRE-motifs are sufficient to regulate abscisic acid (ABA)-responsive gene expression and might confer in some particular cases, depending on the transcription factors, Ca^{2+}-mediated gene expression (Kaplan *et al.* 2006). Phosphorylation of the ABA-Response Element-Binding protein (AREB) will initiate binding of this transcription factor to the ABRE element. The ABRE elements are often found in promoter sequences of abiotic stress inducible genes (Yamaguchi-Shinozaki *et al.* 2005). Except the ABRE-related sequence, the ABRE G-box and CBF, a DRE/CRT *cis*-regulatory element ((G/A)CCGAC) and the related LTRE-element/CBF-binding site were also uniquely present in the *AtSOS1* promoter. These elements were located 220bp upstream of the ATG translation initiation codon (Table 3). The LTRE contains the same core (CCGAC) as the DRE/CRT element. The DRE/CRT element is involved in both dehydration-responsive and low temperature-responsive gene-expression (Yamaguchi-Shinozaki *et al.* 2005, Mizoi *et al.* 2012), and it is controlled by ABA-independent pathways. Dehydration Responsive Element Binding (DREB) proteins, transcription factors belonging to the APETALA2/Ethylene Responsive Factor (AP2/ERF) family, bind to the DRE/CRT element (Urano *et al.* 2010). These transcription factors can bind minutes after stress onset and activate the transcription of genes. DRE/CRT elements, just as ABRE elements, are often found in abiotic-stress

inducible genes (Mizoi *et al.* 2012). The presence of multiple ABRE-variants and the DRE/CRT element in the *AtSOS1* promoter and the absence of these elements from the *SbSOS1* promoter might indicate that either the AREB and/or DREB transcription factor signaling pathways are possibly responsible for the induction of *AtSOS1* in response to salt treatment.

Except the ABRE-variants and the DRE/CRT elements, the *AtSOS1* promoter had several light inducible *cis*-regulatory elements that were absent from the *SbSOS1* promoter: the RBSc-consensus and the Plastid Response Element (PRE) (Table 3 and Table S2). Also the ABRE G-Box might be involved, besides ABA- and Ca^{2+}-responsive gene expression, in light-inducible gene expression (Lopez-Ochoa *et al.* 2007). In the *SbSOS1* promoter, we identified only one light inducible *cis*-regulatory element that was absent from the *AtSOS1* promoter: Sequence Over-Represented in Light Inducible Promoters (SORLIP) (Table 3 and Table S2). Light inducible *cis*-regulatory elements play a role in light responsiveness. Their role in salt tolerance might be exerted through their influence on the water status by regulation of stomatal opening and closure.

Among the unique *cis*-regulatory elements in the *SbSOS1* promoter were: a MYC-like sequence (CATGTG) and the palindromic MYC-binding site (CACATG) (Table 3 and Table S2), although both elements were located far away, 2202bp, from the ATG translation initiation codon. Both the *SbSOS1*-promoter and the *ATSOS1*-promoter had several copies of the more general MYC-consensus (CANNTG), however most copies were located closer to the ATG translational initiation codon in the *SbSOS1* promoter than in the *ATSOS1* promoter (Table S2). The MYC-like sequence is the core sequence of the NAC-recognition site, which can bind transcription factors of the NAC family: the MYC proteins. These abiotic Stress inducible NAC transcription-factors (SNAC) are ABA dependent and are part of the early response to dehydration stress (Yamaguchi-Shinozaki *et al.* 2005, Nakashima *et al.* 2012). A MYB-core binding site (CNGTT(A/G)) and a MYB-recognition site ((A/T)AACCA) were present in both the *SbSOS1* promoter and the *AtSOS1* promoter. The MYB-core binding and recognition site are bound by MYB proteins, also part of the SNAC family of transcription factors (Yamaguchi-Shinozaki *et al.* 2005, Nakashima *et al.* 2012). MYB and MYC proteins are transcriptional activators in the ABA-dependent osmotic stress pathway and are part of the slow- and adaptive response to osmotic stress (Yamaguchi-Shinozaki et al 2005, Nakashima *et al.* 2012).

Additional *cis*-regulatory elements uniquely present in the *SbSOS1* promoters and absent from the *AtSOS1* promoter were, among others, elements that are involved in auxin

responsiveness: Small Auxin Up RNA (SAUR), SUlphur Responsive Element (SuRE) and Activation Sequence Factor-1 (Asf-1) (Table 3 and Table S2). Auxin promotes cell elongation (Davies 2010). How auxin is involved in gene regulation in response to NaCl remains to be established. Another *cis*-regulatory element, which was uniquely present in the *SbSOS1* promoter was the TAAAG-motif (Table 3 and Table S2). The TAAAG-motif is a target site for Dof1 transcription-factors (Plesch *et al.* 2001). Dof1 transcription-factors promote nitrogen assimilation and induces transcription of genes involved in the carbon skeleton production (Yanagisawa *et al.* 2004), but they might also be involved in guard cell-specific gene expression (Plesch *et al.* 2002). A *cis*-regulatory element substantially more, but not uniquely, present in *SbSOS1* promoter was the ARR1-binding element (Table 3 and Table S2). The ARR1-binding element binds Response Regulators (ARR), which are responsive to cytokinins. Cytokinin is a plant hormone that promotes cell division and differentiation, but also leaf senescence (Davies 2010). ARRs may act as a negative regulator on stress and ABA responsive genes (Tran *et al.* 2010).

The *cis*-regulatory elements containing an ABRE-variant and/or the DRE/CRT element might explain the responsiveness of the *A. thaliana SOS1* promoters in response to salinity, and likewise, the absence of these elements in *Salicornia* might explain the lack of responsiveness to salinity. However, the high-level expression of *SOS1* in *S. dolichostachya* is unlikely to be explained by *cis*-regulatory elements, because constitutive high levels of transcription factors will induce the expression of many more genes besides SOS1 in the same fashion, which renders their induction by environmental stimuli and is unlikely to be beneficial. Perhaps the absence of a negative regulator and/or elevated stability of *SOS1* mRNA in the presence of salt (Shi *et al.* 2003; Chung *et al.* 2008) might explain the high-constitutive *SOS1* mRNA level in *S. dolichostachya.*

5.5. Conclusions

S. dolichostachya and *S. oleracea* had both two copies of *SOS1*, and *S. brachystachya* had one copy. Therefore, copy-number difference cannot explain the previously observed constitutively high expression of *SOS1* in *S. dolichostachya* compared with *S. oleracea* (chapter 4). Identification of putative *cis*-regulatory elements might give indications for the inducibility of the *SOS1* gene in *Arabidopsis.* Transcription factors like AREBs, SNAC and DREBs are inducible by osmotic stress. AREBs and SNAC are both responsive to ABA, and DREBs are part of the fast ABA-independent signaling-pathway. AREBs binds to *cis*-regulatory elements containing an ABRE-variant. These elements were present in higher numbers in the *AtSOS1* promoter than in the *SbSOS1* promoter. Moreover, the DRE/CRT element, which is the binding site for DREBs, was uniquely present in the *AtSOS1* promoter. Binding of AREBs and/or DREBs to the *AtSOS1* promoter might cause the upregulation of *AtSOS1* in response to salt treatment, in addition to regulation by the SOS2-SOS3 complex (Qui *et al.* 2002), but this needs to be experimentally confirmed. The high-level *SOS1* expression of *Salicornia* is not likely to originate form binding of transcription factors to *cis*-regulatory elements. In general, the identification of *cis*-regulatory elements in putative salt tolerance genes could give information about the transcription factors that control these genes.

5.6. Acknowledgements

This work was carried out on the Dutch national e-infrastructure with the support of SURF Foundation.

5.7. Supplementary data

Table S1. Primers used for sequencing of the *Salicornia brachystachya SOS1* promoter.

Table S2. Putative *cis*-acting DNA elements in the *Salicornia brachystachya SOS1* promoter, the partial *Salicornia dolichostachya SOS1* promoter and the *Arabidopsis thaliana SOS1* promoter.

Figure S1. The *Salicornia brachystachya SOS1* promoter sequence.

Figure S2. Alignment of the *Salicornia brachystachya SOS1* promoter, the partial *Salicornia dolichostachya SOS1* promoter and the *Arabidopsis thaliana SOS1* promoter.

5.7. Supplementary data

Table S1. Primers used for sequencing of the *Salicornia brachystachya SOS1* promoter sequence.

Forward / Reverse	Sequence (5'-> 3')
Forward	CGGGCGATGTTTAACAAGTCC
Forward	GGAGGAAGATCCTGCTTGAAC
Forward	CTGATAGGCCTCAACTATCATC
Reverse	CTAGGCACAAAACAGTACGAGC
Reverse	CTCCACCACGCTACTCATCTC

Table S2. Putative *cis*-acting DNA elements in the *Salicornia brachystachya SOS1* promoter sequence (2372), the partial *Salicornia dolichostachya SOS1* promoter sequence (596bp) and the *Arabidopsis thaliana SOS1* promoter sequence (981bp). The corresponding parts of the promoter sequences, based on alignment (Fig. S2), were compared. A green color indicates a higher number of the particular *cis*-regulatory element in *Salicornia* compared with the corresponding sequence in the *AtSOS1* promoter (the darker the green, the bigger the difference), a red color indicates less copies of the *cis*-regulatory element in the *Salicornia SOS1* promoter compared with the *Arabidopsis SOS1* promoter.

***Cis*-acting regulatory element**	**Sequence**	***SbSOS1pr* 1501-2372p**	***SbSOS1pr* 596-1500bp**	***AtSOS1pr* 578-981bp**	***SbSOS1pr* 0-595bp**	***AtSOS1pr* 0-577bp**	***SdSOS1pr* 0-596bp**
-10PEHVPSBD	**TATTCT**	1		1	1		1
-300ELEMENT	**TG(A/T/C)AAA(A/G)(T/G)**	2	2		1	1	1
2SSEEDPROTBANAPA	**CAAACAC**		1		1		1
AACACOREOSGLUB1	**AACAAAC**		2		2		2
ABRELATERD1	**ACGTG**			1	2	2	2
ABRERATCAL	**(C/A)ACG(C/T)G(T/C/G)**					2	
ACGTATERD1	**ACGT**			2	4	2	4
AMYBOX1	**TAACA(A/G)A**	3	2				
ANAERO1CONSENSUS	**AAACAAA**	1	3		2	2	2
ANAERO3CONSENSUS	**TCATCAC**	1					
ARR1AT	**NGATT**	10	9	3	5	3	5
ASF1MOTIFCAMV	**TGACG**		1		1		1
BIHD1OS	**TGTCA**	3	2		1	3	
BOXIINTPATPB	**ATAGAA**	2	1				
CAATBOX1	**CAAT**	11	10	7	6	8	7
CACGTGMOTIF	**CACGTG**					2	
CACTFTPPCA1	**(C/T)ACT**	15	15	3	9	12	9
CANBNNAPA	**CNAACAC**		1		1		1
CARGATCONSENSUS	**CATATG**					2	
CARGCW8GAT	**C(A/T)(A/T)(A/T)(A/T)(A/T)(A/T)(A/T)(A/T)G**	2	2				
CATATGGMSAUR	**CATATG**	2			4		4
CBFHV	**(G/A)(C/T)CGAC**					1	
CCAATBOX	**CCAAT**	4	1	2		1	
CEREGLUCBOX2PSLEGA	**TGAAAACT**		2				
CIACADIANLELHC	**CAANNNNATC**	1	1		1	1	1
CPBCSPOR	**TATTAG**	1			1	1	1
CURECORECR	**GTAC**		8		2	6	2
DOFCOREZM	**AAAG**	11	12	2	7	5	7
DPBFCOREDCDC3	**ACACNNG**			1	1	3	2
DRECRTCOREAT	**(G/A)CCGAC**					1	
E2FCONSENSUS	**(A/T)TT(C/G)(C/G)C(C/G)(C/G)**	1					
EBOXBNNAPA	**CANNTG**	4	2	6	6	2	6
EECCRCAH1	**GACTTNC**	1	2				
GAREAT	**TAACAA(A/G)**	4	3				

Cis-acting regulatory element	Sequence	*SbSOS1pr* 1501-2372bp	*SbSOS1pr* 596-1500bp	*AtSOS1pr* 578-981bp	*SbSOS1pr* 0-595bp	*AtSOS1pr* 0-577bp	*SdSOS1pr* 0 - 596bp
GATABOX	GATA	4	9	5	7	6	7
GT1CONSENSUS	G(A/G)(A/T)AA(A/T)	7	23	4	5	10	5
GT1CORE	GGTTAA		1		1		
GT1GMSCAM4	GAAAAA	2	9		2	5	2
GT1MOTIFPSRBCS	(T/G)(A/T)GTG(A/G)(A/T)AA(A/T)(A/G)(A/T)		1				
GTGANTG10	GTGA	3	3	2	5	6	5
HEXMOTIFTAH3H4	ACGTCA				1		1
IBOX	GATAAG	2			1	1	1
IBOXCORE	GATAA	3	3	1	2	4	2
IBOXCORENT	GATAAG(A/G)	1				1	
INRNTPSADB	(C/T)TCANT(C/T)(C/T)	4	3	2	2	1	2
L1BOXATPDF1	TAAATG(A/T)		1	1			
LECPLEACS2	TAAAATAT	1			1		1
LTRECOREATCOR15	CCGAC					1	
MARABOX1	AATAAA(C/T)AAA		1				
MARTBOX	TT(A/T)T(T/W)TT(T/W)TT	4	3		10	3	1
MYB1AT	(A/T)AACCA		1		1	2	1
MYBCORE	CNGTT(A/G)	2	1		1	1	
MYBCOREEATCYCB1	AACGG		1				
MYBGAHV	TAACAAA	3	2				
MYBPZM	CC(A/T)ACC	1		1			
MYBST1	GGATA	1					
MYCATERD1	CATGTG	1					
MYCATRD22	CACATG	1					
MYCCONSENSUSAT	CANNTG	4	2	6	6	2	6
NODCON1GM	CTCTT				1		1
NODCON2GM	CTCTT	1	2		4	3	5
NTBBF1ARROLB	ACTTTA	2	1				
OSE1ROOTNODULE	AAAGAT				1		1
OSE2ROOTNODULE	CTCTT	1	2		4	3	5
POLASIG1	AATAAA	1	6	1		5	
POLASIG2	AATTAAA		2		1	3	2
POLASIG3	AATAAT	4	2		3	1	3
POLLEN1LELAT52	AGAAA	7	5		2	4	2
PREATPRODH	ACTCAT	1	1				
PRECONSCRHSP70A	(C/G)CGA(C/T)N(A/G)NNNNNNNNNNNNNNN(A/T/C)(A/T/G)					1	

Cis-acting regulatory element	Sequence	*SbSOS1pr* 1501-2372bp	*SbSOS1pr* 596-1500bp	*AtSOS1pr* 578-981bp	*SbSOS1pr* 0-595bp	*AtSOS1pr* 0-577bp	*SdSOS1pr* 0 - 596bp
PYRIMIDINEBOXHVEPB1	TTTTTTCC		2		1		1
PYRIMIDINEBOXOSRAMY1A	CCTTTT		1		1	1	1
RAV1AAT	CAACA	4	2	2	4	3	4
RBCSCONCENSUS	AATCCAA			1		1	
REALPHALGLHCB21	AACCAA		1		1	1	1
REBETALGLHCB21	CGGATA	1					
RHERPATEXPA7	(T/G)CACG(A/T)			1	1		1
ROOTMOTIFTAPOX1	ATATT	8	8	4	5	12	1
RYREPEATBNNAPA	CATGCA			1			
RYREPEATLEGUMINBOX	CATGCA(C/T)			1			
S1FBOXSORPS1L21	ATGGTA	1				1	
SEF1MOTIF	AACCCA					1	
SEF4MOTIFGM7S	ATATTTA(A/T)(A/T)	4	1	1	3	1	3
SITEIIATCYTC	TGGGC(C/T)	1					
SORLIP1AT	GCCAC	1			2		2
SORLIP2AT	GGGCC	2					
SP8BFIBSP8BIB	TACTATT				1	1	1
SURECOREATSULTR11	GAGAC	1			1		2
T/GBOXATPIN2	AACGTG			1			
TAAAGSTKST1	TAAAG	3	3		1		1
TATABOX2	TATAAAT		1	2		1	
TATABOX3	TATTAAT	1			2	1	2
TATABOX4	TATATAA				1		1
TATABOX5	TTATTT	5	5		1	4	1
TATABOXOSPAL	TATTTAA	1					
TBOXATGAPB	ACTTTG	1					
TGACGTVMAMY	TGACGT				1		1
TRANSINITDICOTS	A(A/C)NAUGGC	1			1		1
TRANSINITMONOCOTS	(A/G)(A/C)NAUGGC	1			1		1
UP2ATMSD	AAACCCTA		1				
UPRMOTIFIIAT	CCNNNNNNNNNNNNCCACG	1					
WBOXATNPR1	TTGAC	1	2		1	1	1
WBOXHVISO1	TGACT		1	1	3	1	3
WBOXNTCHN48	TTGAC			1		1	
WBOXTERF3	TGAC(C/T)		2	1	3	2	3
WRKY71OS	TGAC	3	5	1	5	5	4
WUSATAg	TTAATGG					1	
XYLAT	ACAAAGAA	1	1				

Figure S1. The *Salicornia brachystachya SOS1* promoter sequence.

CGATTTTAAGTAATATGCGGGCGATGTTTAACAAGTCCCCATATTAATTGCAACGATTTGACATTTTTTATGAAATGAATGAT
AAGGCGTGGTGGGGCATGGTAGGGGGGAAATTTAAAGGAATGCCAGACTGTCATCACTAATAAGGCAAATAAGACAAAGGATG
TGGCATGTGAAATGGACTCGTTTGATCATTATTATTCTTTGTTGTTGCTGATCAATCTGTTGCAGCCAGCCATCTTAGTAACA
TAGTGAAAGTATCCGATCTCAAACCGATCTGATCTGTAGGAATATATAGTAGTCCTTTCATTTTTCTACTTTACTTATTACTC
CCTCTGTCCCAAATTACATGCTACAATTTTCTTATTTGGATGTTCTCAAATACTTGCTATATTTCTTTTTTATAGTCTATGGG
CCCCACATATATTTTACATTTTTACCCTCTTTTTTTGTCATTTTTATCTTATTTAAGCCAATCAATTTCTCCCCCTTAACTTT
ACCAATGGTGGACCCACATTAAAATTGATTTGCATAAAAATTTGTTCCAAAATAGAATGTAGCATGTAAAATGAGACGGAGGG
AGTATTATTATTGATTTATTGGAGCACCAAGATCTGTTGTTTTTCTTATTTCTTATGCATATGATTTGATGAATTGGACTCGA
GTGGTCCTGTTCTGTTTCTTGAACTTATCATTTTTGTTTTTTTTAATATTATGAATTTTTTTAGACCTTATTTTTTTGTTACA
CGATTTTGTTATAAGTAGTAATTAGATTCTATTTGAAACTTTTTTTTGGAGGAAGATCCTGCTTGAACTTTGTTAATATAAAAA
AAAAAAACTCTCCACCACGCTACTCATCTCAATCTCCATCTTAGGCTCTGTTCTTTTGGACTTAACCACCTTTATCATTCATA
ATATACGAAAAAACTAAAATAAAAAGAGCAAACCTAAACCCTAGGAGTTTGTTATAATTTCACTTGTGCTCTGTTTATTTTGA
GTGAAAACTATTTTCAATGGAGATTGTCAAACACATTTTTTCCTTTGGTTGTATGACCATCGAAAAATTTAAGGGGAAAAAGT
AGAAGCAAAAAAGGTACGGATGGAAGTACAGAACAAAATATGGAGGCAAAATGAAAATAGAAGAAAAATAGGGTGGAAGGATG
ATATATATAGTAGATAACAAAACATAATGAGTAAGATGAAATATGCGGCATACTAGAGAACGAAAATGATTTTCCTGTTCAGA
AGATTTGTTTTTTTTCACTCTTTAGAAATTCATGTTCTTCCACTAAGCAAAACAAACAAAGAAACATGACTAGGGGGTGGAAA
ATATTTTCATTTTATTTGGTCCATATCTATTTATTTATTAAAAGCTGAAAATGGATTTTTTTTTTCTTTTAAGTTTTGATTTG
AAGAACCACCTATGTAAATGACAACTTTATCATAACCTTGCCCCCTAAAAAACATGATTATTGTACATTCATAAATATCAATT
TGATTTTTTTAATATGATTTTTTTTTTTCCAATTTCTAACAAGCAATATGGTGCTCAAATAAATTAGAGAGAACATAATAGTT
TTCAGCAATCATATACTTCAACCGTTTTTAATTACTCCGTAGATACTACACTTTTCATGTTGACGCCATATACAATTATTCAT
TGTCTATTTTGCACACAATTAAATGCAACATTAGACACAAATTCTGATAGGCCTCAACTATCATCTTACTCTATAAATTCGTT
GCGAAAAAAAAACTAGGCACAAAACAGTACGAGCCTCACGTAAAAATCAATACTGCATATGAAATAGTCATAAGTATAATGTTG
ATGCCATATACAATTATTCATTTCGTCTATTTTGTAGACACTTAAAAATAACATTAGACACAAATTCTGATAAAATATATAAT
ATATTTCTAAGTTTCAACTATCAACTTACCGTATCAAATTCCTTGAAAAAAAAACACTAGACACACAACAGCAGGAGTCTCAC
ATAAAAATAGTAATGCATATGAAATAGTCACAACTATAATTTAAATTTTTTTAACCAAAAAAAAAGTCCCACTGTCAAAAAAA
AAAAAAAAAAAATCAGGCACAAAAGTAGTTTTTTTTGGTAAAAGGTGACGTGTGGCTCTTGAGAACTTATATTAATAGTTGTA
CACCTATCATCTGGCTCTCTCATCCCTTGAAAACAAACACAAACAAACAACTAATGATGATACATTGATGATTGATAAGCATC
TATATCTATATTACTCCCTAAATCTTGTTGACTCCTCTTTAATTTTGCAACTTCATCTTTTTTTCCTCACTTTCTCTTTCATC
TAATAATCAATTATTCTCTCTTCCGCCATTGTTGCAGCCACGGCAGCC**ATG**

Figure S2. Alignment of the *Salicornia brachystachya SOS1* promoter sequence (2372), the partial *Salicornia dolichostachya SOS1* promoter sequence (596bp) and the *Arabidopsis thaliana SOS1* promoter sequence (981bp). Alignment was performed with CLUSTALW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/).

```
SdSOS1pr_596bp      ------------------------------------------------------------
SbSOS1pr_2372bp     CGATTTTAAGTAATATGCGGGCGATGTTTAACAAGTCCCCATATTAATTGCAACGATTTG
AtSOS1pr_981bp      ------------------------------------------------------------

SdSOS1pr_596bp      ------------------------------------------------------------
SbSOS1pr_2372bp     ACATTTTTTATGAAATGAATGATAAGGCGTGGTGGGGCATGGTAGGGGGGAAATTTAAAG
AtSOS1pr_981bp      ------------------------------------------------------------

SdSOS1pr_596bp      ------------------------------------------------------------
SbSOS1pr_2372bp     GAATGCCAGACTGTCATCACTAATAAGGCAAATAAGACAAAGGATGTGGCATGTGAAATG
AtSOS1pr_981bp      ------------------------------------------------------------

SdSOS1pr_596bp      ------------------------------------------------------------
SbSOS1pr_2372bp     GACTCGTTTGATCATTATTATTCTTTGTTGTTGCTGATCAATCTGTTGCAGCCAGCCATC
AtSOS1pr_981bp      ------------------------------------------------------------

SdSOS1pr_596bp      ------------------------------------------------------------
SbSOS1pr_2372bp     TTAGTAACATAGTGAAAGTATCCGATCTCAAACCGATCTGATCTGTAGGAATATATAGTA
AtSOS1pr_981bp      ------------------------------------------------------------

SdSOS1pr_596bp      ------------------------------------------------------------
SbSOS1pr_2372bp     GTCCTTTCATTTTTCTACTTTACTTATTACTCCCTCTGTCCCAAATTACATGCTACAATT
AtSOS1pr_981bp      ------------------------------------------------------------

SdSOS1pr_596bp      ------------------------------------------------------------
SbSOS1pr_2372bp     TTCTTATTTGGATGTTCTCAAATACTTGCTATATTTCTTTTTTATAGTCTATGGGCCCCA
AtSOS1pr_981bp      ------------------------------------------------------------

SdSOS1pr_596bp      ------------------------------------------------------------
SbSOS1pr_2372bp     CATATATTTTACATTTTTACCCTCTTTTTTTGTCATTTTTATCTTATTTAAGCCAATCAA
AtSOS1pr_981bp      ------------------------------------------------------------

SdSOS1pr_596bp      ------------------------------------------------------------
SbSOS1pr_2372bp     TTTCTCCCCCTTAACTTTACCAATGGTGGACCCACATTAAAATTGATTTGCATAAAAATT
AtSOS1pr_981bp      ------------------------------------------------------------

SdSOS1pr_596bp      ------------------------------------------------------------
SbSOS1pr_2372bp     TGTTCCAAAATAGAATGTAGCATGTAAAATGAGACGGAGGGAGTATTATTATTGATTTAT
AtSOS1pr_981bp      ------------------------------------------------------------

SdSOS1pr_596bp      ------------------------------------------------------------
SbSOS1pr_2372bp     TGGAGCACCAAGATCTGTTGTTTTTCTTATTTCTTATGCATATGATTTGATGAATTGGAC
AtSOS1pr_981bp      ------------------------------------------------------------

SdSOS1pr_596bp      ------------------------------------------------------------
SbSOS1pr_2372bp     TCGAGTGGTCCTGTTCTGTTTCTTGAACTTATCATTTTTGTTTTTTTTTAATATTATGAAT
AtSOS1pr_981bp      ------------------------------------------------------------

SdSOS1pr_596bp      ------------------------------------------------------------
SbSOS1pr_2372bp     TTTTTTAGACCTTATTTTTTTGTTACACGATTTTGTTATAAGTAGTAATTAGATTCTATT
AtSOS1pr_981bp      ------------------------------------------------------------

SdSOS1pr_596bp      ------------------------------------------------------------
SbSOS1pr_2372bp     TGAAACTTTTTTTGGAGGAAGATCCTGCTTGAACTTTGTTAATATAAAAAAAAAAAACTC
AtSOS1pr_981bp      ------------------------------------------------------------



SdSOS1pr_596bp      ------------------------------|-----------------------------
SbSOS1pr_2372bp     TCCACCACGCTACTCATCTCAATCTCCATCT|AGGCTCTGTTCTTTTGGACTTAACCACC
AtSOS1pr_981bp      ------------------------------|AGGATCG--ACGGTTGGTCG---------

SdSOS1pr_596bp      ------------------------------------------------------------
```

```
SbSOS1pr_2372bp     TTTATCATTCATAATATACGAAAAAACTAAAATAAAAAGAGCAAACCTAAACCCTAGGAG
AtSOS1pr_981bp      ------------------------------ATTAATCAATACTAACCCAAACAG------

SdSOS1pr_596bp      ------------------------------------------------------------
SbSOS1pr_2372bp     TTTGTTATAATTTCACTTGTGCTCTGTTTATTTTGAGTGAAAACTATTTTCAATGGAGAT
AtSOS1pr_981bp      -------TCTTTTAACATG-----------------------------------------

SdSOS1pr_596bp      ------------------------------------------------------------
SbSOS1pr_2372bp     TGTCAAACACATTTTTTCCTTTGGTTGTATGACCATCGAAAAATTTAAGGGGAAAAAGTA
AtSOS1pr_981bp      -------CACACTTGATACCTTG-------------------------------------

SdSOS1pr_596bp      ------------------------------------------------------------
SbSOS1pr_2372bp     GAAGCAAAAAAGGTACGGATGGAAGTACAGAACAAAATATGGAGGCAAAATGAAAATAGA
AtSOS1pr_981bp      ---------TAGGTGCTTCTG----------------------GGTAGCATGAAAATG--

SdSOS1pr_596bp      ------------------------------------------------------------
SbSOS1pr_2372bp     AGAAAAATAGGGTGGAAGGATGATATATATAGTAGATAACAAAACATAATGAGTAAGATG
AtSOS1pr_981bp      ------------------------------------------------------------A

SdSOS1pr_596bp      ------------------------------------------------------------
SbSOS1pr_2372bp     AAATATGCGGCATACTAGAGAACGAAAATGATTTTCCTGTTCAGAAGATTTGTTTTTTTT
AtSOS1pr_981bp      TAATTGGTGTTATTCTGATATGCGAAATTGAAAATGTTG---------------------

SdSOS1pr_596bp      ------------------------------------------------------------
SbSOS1pr_2372bp     CACTCTTTAGAAATTCATGTTCTTCCACTAAGCAAAACAAACAAAGAAACATGACTAGGG
AtSOS1pr_981bp      ----------------TTGTGCATCAGCTAAGCATAATTGAAAGT---------------

SdSOS1pr_596bp      ------------------------------------------------------------
SbSOS1pr_2372bp     GGTGGAAAATATTTTCATTTTATTTGGTCCATATCTATTTATTTATTAAAAGCTGAAAAT
AtSOS1pr_981bp      -TTAAAAAATATTATAAATTTAACG-----------------------------------

SdSOS1pr_596bp      ------------------------------------------------------------
SbSOS1pr_2372bp     GGATTTTTTTTTTCTTTTAAGTTTTGATTTGAAGAACCACCTATGTAAATGACAACTTTA
AtSOS1pr_981bp      -------------------------------TGAAG----------------------

SdSOS1pr_596bp      ------------------------------------------------------------
SbSOS1pr_2372bp     TCATAACCTTGCCCCCTAAAAAACATGATTATTGTACATTCATAAATATCAATTTGATTT
AtSOS1pr_981bp      -------------------------------------AATTGAAAGGTATCAGATTAATGT

SdSOS1pr_596bp      ------------------------------------------------------------
SbSOS1pr_2372bp     TTTTAATATGATTTTTTTTTTTCCAATTTCTAACAAGCAATATGGTGCTCAAATAAATTA
AtSOS1pr_981bp      TTTGGATCAG------------CCGAATAGAGTCAG----------GCTAATCCAAATT-

SdSOS1pr_596bp      ------------------------------------------------------------
SbSOS1pr_2372bp     GAGAGAACATAATAGTTTTCAGCAATCATATACTTCAACCGTTTTTAATTACTCCGTAGA
AtSOS1pr_981bp      -------CTCAAATATTTTCG----------------------------AAACCTTAAA

SdSOS1pr_596bp      ------------------------------------------------------------
SbSOS1pr_2372bp     TACTACACTTTTCATGTTGACGCCATATACAATTATTCATTGTCTATTTTGCACACAATT
AtSOS1pr_981bp      CACATCAAACTCCACGAAGTAAACACACTCACA------------------CACATATCA

SdSOS1pr_596bp      ------------------------------------------------------------
SbSOS1pr_2372bp     AAATGCAACATTAGACACAAATTCTGATAGGCCTCAACTATCATCTTACTCTATAAATTC
AtSOS1pr_981bp      AAATCCACATTATAAATGTATTTTTGGTAGGCTAG-----TCGTCTAATTCTAAAAT---



SdSOS1pr_596bp      -------------------------------------CTCACGTAAAAATCAATACTG---
SbSOS1pr_2372bp     GTTGCGAAAAAAAAACTAGGCACAAAACAGTACGAGCCTCACGTAAAAATCAATACTG---
AtSOS1pr_981bp      --AGCCAATTTACATTTG---CAATTGTTTAT---TCAAAAAAGTACAACAGTAAGAGGT
                                                     *:.*..:.:*.*:**.**. .

SdSOS1pr_596bp      CATATGAAATAG-TCATAAGTATAATGT-TGATGCCATATACAATTATTCATTTCGTCTA
SbSOS1pr_2372bp     CATATGAAATAG-TCATAAGTATAATGT-TGATGCCATATACAATTATTCATTTCGTCTA
AtSOS1pr_981bp      CTTATCAATTAAATCATAAGAAAAATATGTCCCAAAATGTCAAAAAACCAACTACGAATA
                    *:*** **:**. *******:*:***.* * . ...**.*..**::*  .* *:**:.**
```

```
SdSOS1pr_596bp      TTTTGTAGACACTTAAAAATAACATTAG-ACACAAATTCTGATAAAATATATAATATATT
SbSOS1pr_2372bp     TTTTGTAGACACTTAAAAATAACATTAG-ACACAAATTCTGATAAAATATATAATATATT
AtSOS1pr_981bp      TTTCTTTTCCTTTTAAGGCCAAACCAGGTTTTGGAATACTTATTACAAATAAACTCAATT
                    ***  *: .*: ****... **.. :.* : : .***:** **:*.*:***:*.*.:***

SdSOS1pr_596bp      TCTA-AGTCTCAACTATCAACTTACCGTATCAAATTCCTTGAAAAAAAAAAACACTAGACA
SbSOS1pr_2372bp     TCTA-AGTTTCAACTATCAACTTACCGTATCAAATTCCTTG-AAAAAAAAAACACTAGACA
AtSOS1pr_981bp      TCAATAAATTCTTGTACACAAATTAAGCACTAATTTCATTG--TGAAGATACCATAGTCA
                    **:* *.: **:: ** ..*.:*:..* *  **:***.***  :.**.*:**..***:**

SdSOS1pr_596bp      CACAACACTAGGAGTCTCACATAAAAATAGTAATGCATATGAAATAGTCACAACTATAAT
SbSOS1pr_2372bp     CACAACAGCAGGAGTCTCACATAAAAATAGTAATGCATATGAAATAGTCACAACTATAAT
AtSOS1pr_981bp      CATTCACATCTATATGACATCAATCTATA-AAATCCAAAACAAACTATATATTATATTTA
                    ** :...  . .:.* :** .:*:.:*** :*** **:*: *** :.*.:.::.***:::

SdSOS1pr_596bp      TTAAATTTTTTAAACCAAAAAAAAAGAGTCCCATTG---TAAAAAAAATTAAAAAAATCA
SbSOS1pr_2372bp     TTAAATTTTTTTAACCAAAAAAAAA--GTCCCACTGTCAAAAAAAAAAAAAAAAAAATCA
AtSOS1pr_981bp      TATTATTTTTTTGTCATCATTATAT---TTCCATTA----------ATTTATCTAATTCT
                    *:::*******:.:*.:.*::*:*:   * *** *.          *:::*:.:**:**:

SdSOS1pr_596bp      GGCACAAAAGTAGTTTTTTTTGGTAAAAGGTGACGTGTGGCTCTTGAGAACTTATATTAA
SbSOS1pr_2372bp     GGCACAAAAGTAGTTTTTTTTGGTAAAAGGTGACGTGTGGCTCTTGAGAACTTATATTAA
AtSOS1pr_981bp      GGTTCTAATATACTCTTGGTCAGAAAAAATATAAACATTGAAGAATTGGTCGGCTGAAAA
                    ** :*:**:.** * **  * .*:****. : *.. .* *.: :: :*.:*  .*.::**

SdSOS1pr_596bp      TAGTTGTACACCTATCATCTGGCTCTCTCATCCCTTGAAAACAAACACAAACAAACAACT
SbSOS1pr_2372bp     TAGTTGTACACCTATCATCTGGCTCTCTCATCCCTTGAAAACAAACACAAACAAACAACT
AtSOS1pr_981bp      TTGTG--AAAAATATATAGCAGAAAAATATGATAATGTTATCATAAACAAAATTAATAGT
                    *:**   *.*..***.::  .*.:.:.*.: . .:**::*:**:*.*****.::*.:* *

SdSOS1pr_596bp      AATGATGATACATTGATGATTGATAAGCATCTATATCTATATTACTCCCTAAATCTTGTT
SbSOS1pr_2372bp     AATGATGATACATTGATGATTGATAAGCATCTATATCTATATTACTCCCTAAATCTTGTT
AtSOS1pr_981bp      AAAATTTAATTTTAATTTAACACTACAGTACTATACACGTGTGTATGTATAGCTCTATAA
                    **:.:* *:: :*:.:* *: ..**.. ::***** . .*.* :.*  .**..***: ::

SdSOS1pr_596bp      GACTCCTCTTTAATTTTGCAACTTCATCTTTTTTTCCTCACTTTCTCTTTCATCTAATAA
SbSOS1pr_2372bp     GACTCCTCTTTAATTTTGCAACTTCATCTTTTTTTCCTCACTTTCTCTTTCATCTAATAA
AtSOS1pr_981bp      GTATTTACTCTCTTTCAGCTATTTATTTTTTCAGTGAACGAG----CATTCTTCTTCTTC
                    *:.*  :** *.:** :**:* **.:* *** : * .:*..     *:***:***:.*:.

SdSOS1pr_596bp      TCAATTATTCTCTCTTCCGCCATTGTTGCAGCCACGGCAGCCATG
SbSOS1pr_2372bp     TCAATTATTCTCTCTTCCGCCATTGTTGCAGCCACGGCAGCCATG
AtSOS1pr_981bp      CTCTGTGTTGTTGCTTCTTAGATATATTCAAATAA------AATG
                      .: *.** *  ****  . **: :* **.. *.      .***
```